

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2002, 22:26:45 ; Search time 8826.93 Seconds
(without alignments)
11851.432 Million cell updates/sec

Title: US-09-645-593-8

Perfect score: 4999

Sequence: 1 ctcaagcatacggaagg.....tcnattgtanagnccttg 4999

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: gb.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description

1	4399	100.0	4999	6	AX089358	Sequence	8 from Patent WO0116340.	4999 bp	DNA	linear	PAT 21-MAR-2001
2	25	0.5	26541	2	AC095913	Rattus no					
3	25	0.5	113725	2	AC108615	Rattus no					
c	4	0.5	138733	9	AL136233	Human DNA					
5	24	0.5	32	6	AX089372	Sequence					
6	24	0.5	1683	8	AF240004	Sesamum i					
7	24	0.5	133108	2	AC034196	Homo sapi					
8	24	0.5	148598	2	AC022096	Homo sapi					
c	9	24	0.5	152602	2	AC026193	Homo sapi				
10	24	0.5	162433	2	AC024377	Homo sapi					
c	11	24	0.5	163272	2	AC018496	Homo sapi				
12	24	0.5	165077	2	AC026190	Homo sapi					
c	13	24	0.5	169770	9	AC027763	Homo sapi				
14	24	0.5	173878	2	AC040977	Homo sapi					
c	15	24	0.5	176217	9	AC034192	Homo sapi				
16	24	0.5	176219	9	AC087090	Homo sapi					
17	24	0.5	191050	2	AC073565	Mus muscu					
18	24	0.5	202704	10	AC090843	Mus muscu					
c	19	23	0.5	29	6	AX089374	Sequence				
20	23	0.5	420	5	AF013235	Quiscatus					
c	21	23	0.5	606	5	AB067610	Ophthalmo				
22	23	0.5	804	3	DDITC4RAN	L09720 Dictyosteli					
23	23	0.5	1738	3	DMG43737	X78119 P. amygdalus					
24	23	0.5	1905	8	PABTPR1	AF466307 Dictyoste					
25	23	0.5	4041	3	AF466307	Rana rugo					
c	26	23	0.5	4644	5	AB017353	Homo sapi				
27	23	0.5	42265	9	AB010504	Homo sapi					
c	28	23	0.5	67687	2	AC101305	Mus muscu				
29	23	0.5	72778	2	AC107045	Mus muscu					
30	23	0.5	75572	2	AC101138	Mus muscu					
c	31	23	0.5	77538	9	AC004500	Homo sapi				
32	23	0.5	97057	9	AL627344	Human DNA					
c	33	23	0.5	113725	2	AC108615	Rattus no				
34	23	0.5	126910	9	AC026794	Homo sapi					
c	35	23	0.5	128151	9	AP001966	Homo sapi				
36	23	0.5	134757	9	AC007870	Genomic s					
c	37	23	0.5	137336	8	AC007915	Genomic s				
38	23	0.5	138752	9	AL161627	Human DNA					
c	39	23	0.5	143799	9	AC011059	Homo sapi				
40	23	0.5	147225	2	AC022498	Homo sapi					
c	41	23	0.5	148598	9	HSBA51C14	Human DNA				
42	23	0.5	149133	2	AC079089	Homo sapi					
c	43	23	0.5	150159	2	AC018791	Homo sapi				
44	23	0.5	151301	2	AL354717	Homo sapi					
45	23	0.5	154769	2	AL671890	Mus muscu					

ALIG MENTS

RESULT 1

AX089358

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

flax.

Linum usitatissimum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Malpighiales; Linaceae; Linum.

1 (bases 1 to 4999)

Chaudhary, S., van Rooijen, G., Moloney, M.M. and Singh, S.

Flax seed specific promoters

PATENT: WO 0116340-A 8 08-MAR-2001;

SemioSys Genetics Inc. (CA); COMMONWEALTH SCIENTIFIC AND

INDUSTRIAL RESEARCH ORGANISATION (AU)

LOCATION/Qualifiers

1..4999

/organism="Linum usitatissimum"

/db_xref="taxon:4006"

BASE COUNT 1340 a 1125 c 1192 g 1238 t 104 others


```

* 3849 5491: contig of 1643 bp in length
* 5492 5591: gap of unknown length
* 5592 6820: contig of 1229 bp in length
* 6821 9295: gap of unknown length
* 9296 9395: contig of 2375 bp in length
* 9396 10779: gap of unknown length
* 10780 11955: contig of 1384 bp in length
* 11956 12058: gap of unknown length
* 12059 13529: contig of 1076 bp in length
* 13530 13629: gap of unknown length
* 13630 14871: contig of 1473 bp in length
* 14872 14970: gap of unknown length
* 14971 16578: contig of 1242 bp in length
* 16579 17297: contig of 1607 bp in length
* 17298 18028: gap of unknown length
* 18029 19318: contig of 1250 bp in length
* 19319 20918: contig of 1190 bp in length
* 20919 21018: gap of unknown length
* 21019 22505: contig of 1601 bp in length
* 22506 22605: gap of unknown length
* 22606 24325: contig of 1487 bp in length
* 24326 24426: gap of unknown length
* 24427 25430: contig of 1720 bp in length
* 25431 25530: gap of unknown length
* 25531 26541: contig of 1004 bp in length
* 26542 26541: contig of 1012 bp in length.

```

FEATURES

```

Location/Qualifiers
1..26541
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/cd_xref="CH230-10M9"

```

```

BASE COUNT 6642 a 5472 c 5923 g 6839 t 1665 others
ORIGIN

```

```

Query Match 0.5% Score 25; DB 2; Length 26541;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3808 atatatgtaataataataataa 3832
|||||
Db 16191 ATATATGTAATAATAATAATAA 16215
|||||

```

```

RESULT 3
AC108615 113725 bp DNA linear HTG 31-JAN-2002
LOCUS Rattus norvegicus clone CH230-290E9, *** SEQUENCING IN PROGRESS
DEFINITION *** 62 unordered pieces.
AC108615
AC108615.1 GI:18449450
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 113725)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbora,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouch,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

```

REFERENCE

AUTHORS

```

* 3974 4073: contig of 3973 bp in length
* 4074 7436: gap of unknown length
* 7437 7537: gap of unknown length
* 7538 10805: contig of 3269 bp in length
* 10806 10905: gap of unknown length
* 10906 13707: contig of 2802 bp in length

```

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,K., Gorielli,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonaik,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Sutton,A., Svatek,A., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Tansey,J., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wall,R., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

TITLE

JOURNAL

* 13708 13807: gap of unknown length
 * 13808 16732: contig of 2925 bp in length
 * 16833 16832: gap of unknown length
 * 16833 19777: contig of 2145 bp in length
 * 18978 19077: gap of unknown length
 * 19078 21825: contig of 2748 bp in length
 * 21826 21925: gap of unknown length
 * 21926 24091: contig of 2166 bp in length
 * 24092 24191: gap of unknown length
 * 24192 25964: contig of 1773 bp in length
 * 25965 26064: gap of unknown length
 * 26065 28909: contig of 2845 bp in length
 * 28910 29009: gap of unknown length
 * 30010 30413: contig of 1404 bp in length
 * 30414 30513: gap of unknown length
 * 30514 31753: contig of 1240 bp in length
 * 31754 31853: gap of unknown length
 * 31854 34448: contig of 2595 bp in length
 * 34449 34548: gap of unknown length
 * 34549 36351: contig of 1803 bp in length
 * 36352 38451: gap of unknown length
 * 36452 38532: contig of 2081 bp in length
 * 38533 38632: gap of unknown length
 * 38633 40773: contig of 2141 bp in length
 * 40774 40873: gap of unknown length
 * 40874 42725: contig of 1852 bp in length
 * 42726 42825: gap of unknown length
 * 42826 44501: contig of 1878 bp in length
 * 44502 44601: gap of unknown length
 * 44602 46744: contig of 2143 bp in length
 * 46745 46844: gap of unknown length
 * 46845 48827: contig of 1883 bp in length
 * 48828 51179: contig of 2352 bp in length
 * 51180 51279: gap of unknown length
 * 51280 52478: contig of 1199 bp in length
 * 52479 52578: gap of unknown length
 * 52579 54000: contig of 1422 bp in length
 * 54001 54100: gap of unknown length
 * 54101 55860: contig of 1760 bp in length
 * 55861 55960: gap of unknown length
 * 55961 58011: contig of 2051 bp in length
 * 58012 58111: gap of unknown length
 * 58112 59783: contig of 1672 bp in length
 * 59784 59883: gap of unknown length
 * 59884 61804: contig of 1921 bp in length
 * 61805 61904: gap of unknown length
 * 61905 63014: contig of 1110 bp in length
 * 63015 63114: gap of unknown length
 * 63115 64428: contig of 1314 bp in length
 * 64429 64528: gap of unknown length
 * 64529 65848: contig of 1320 bp in length
 * 65849 67951: contig of 2003 bp in length
 * 67952 68051: gap of unknown length
 * 68052 69813: contig of 1762 bp in length
 * 69814 69913: gap of unknown length
 * 69914 71245: contig of 1332 bp in length
 * 71246 71346: gap of unknown length
 * 71346 72906: contig of 1561 bp in length
 * 72907 73006: gap of unknown length
 * 73007 74675: contig of 1669 bp in length
 * 74676 74775: gap of unknown length
 * 74776 76188: contig of 1413 bp in length
 * 76189 76288: gap of unknown length
 * 76289 77815: contig of 1527 bp in length
 * 77816 77915: gap of unknown length
 * 77916 79872: contig of 1957 bp in length
 * 79873 79972: gap of unknown length
 * 79973 81380: contig of 1408 bp in length
 * 81381 81480: gap of unknown length
 * 81481 82512: contig of 1032 bp in length
 * 82513 82612: gap of unknown length

* 82613 84502: contig of 1890 bp in length
 * 84503 84602: gap of unknown length
 * 84603 85955: contig of 1353 bp in length
 * 85956 86055: gap of unknown length
 * 86056 87130: contig of 1075 bp in length
 * 87131 87230: gap of unknown length
 * 87231 88779: contig of 1549 bp in length
 * 88780 90359: gap of unknown length
 * 90360 90459: gap of unknown length
 * 90460 91498: contig of 1039 bp in length
 * 91499 91598: gap of unknown length
 * 91599 93463: contig of 1865 bp in length
 * 93464 93563: gap of unknown length
 * 93564 94976: contig of 1413 bp in length
 * 94977 95076: gap of unknown length
 * 95077 96234: contig of 1158 bp in length
 * 96235 96334: gap of unknown length
 * 96335 97663: contig of 1329 bp in length
 * 97664 97763: gap of unknown length
 * 97764 99250: contig of 1487 bp in length
 * 99251 99350: gap of unknown length
 * 99351 100809: contig of 1459 bp in length
 * 100810 100909: gap of unknown length
 * 100910 101932: contig of 1023 bp in length
 * 101933 102032: gap of unknown length
 * 102033 103042: contig of 1010 bp in length
 * 103043 103142: gap of unknown length
 * 103143 104157: contig of 1015 bp in length
 * 104158 104257: gap of unknown length

Query Match 0.5%; Score 25; DB 2; Length 113725;

Best Local Similarity 100.0%; Pred.No. 0.11;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3808 atatatgaataataataataa 3832

Db 13816 ATATATGTAATAATAATAATAA 13840

RESULT 4

AL1362233/C
 LOCUS AL1362233 138733 bp DNA linear PRI 21-OCT-2001
 DEFINITION Human DNA sequence from clone RP11-9E13 on chromosome 10, complete sequence.
 ACCESSION AL1362233
 VERSION AL1362233.14 GI:16408601
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 Howden, P.
 Direct Submission
 Submitted (20-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Oct 24, 2001 this sequence version replaced gi:15131189.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

COMMENT

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given

chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10>
RP11-9E13 is from the library RPCT-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

IMPORTANT: This sequence is not the entire insert of clone RP11-9E13. It may be shorter because we sequence overlapping sections only once, except for a short overlap. The true right end of clone RP11-9E13 is at 138733 in this sequence. The true right end of clone RP11-153K11 is at 200000 in this sequence.

FEATURES	Location/Qualifiers
source	1. .138733

```
1. .130733
/organism="Homo sapiens"
/db xref="taxon:9606"
```

```
/chromosome="10"  
/clone="RP11-9E13"  
/clone_lib="RP11-11 1"
```

```

/chr10c_110 chr10:110
55963..56014
/note="Sequence from uni-directional dGTP big dye
terminator reads only."

```

62420. 63555
/note="CNC is

```

/notice=cpo 1 and
/evidence=not_experimental
67805. 67836

```

/note="Single clone restriction digest"

```

restriction digest data.
77610..77625
/note="Single clone region Assembly confirmed by

```

restriction di
127610 1280-

```
127010. 120010
/note="CpG island"
/evidence=not experimental
```

BASE COUNT	43280 a	29780 c	28584 g	37089 t
ORIGIN				

Query Match 0.5%; Score 25; DB 9; Length 138733;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25: Conservative 0; Mismatches 0; Indels 0

QY 3815 taataataataataataagccaca 3839

Db 108555 TAATAATAATAATAAGCCACA 108531

RESULT	5
AV000070	

AX063572	LOCUS	AX089372	32 bp	DNA	linear	PAT 21-MAR-2001
	DEFINITION	Sequence 22 from	Patent WO0116340			

ACCESSION	AX089372
VERSION	AX089372.1
GI	13443635

KEYWORDS
SOURCE . synthetic construct.

ORGANISM	synthetic construct	artificial sequence
<i>S. aureus</i>	+	-
<i>E. coli</i>	+	-
<i>B. subtilis</i>	+	-
<i>L. monocytogenes</i>	+	-
<i>C. jejuni</i>	+	-
<i>H. pylori</i>	+	-
<i>M. luteus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. abscessus</i>	+	-
<i>M. chelonae</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-
<i>M. luteus</i>	+	-
<i>M. neoaurum</i>	+	-
<i>M. thermophilus</i>	+	-
<i>M. vacuus</i>	+	-
<i>M. fortuitus</i>	+	-
<i>M. goodii</i>	+	-

REFERENCE 1 (bases 1 to 32)

AUTHORS	Chaudhary, S., van Rooljen, G., Moloney, M.M. and Singh, S.
TITLE	Flax seed specific promoters

JOURNAL
Patent: WO 0116340-A 22 08-MAR-2001;
SemBioSys Genetics Inc. (CA) ; COMMO

FEATURES	INDUSTRIAL RESEARCH ORGANISATION (AU)
Location/Qualifiers	

LEADS	source	LOCATION/QUARTERS
		1. 32

```
/organism="synthetic construct"
/db_xref="taxon:32630"
```

BASE COUNT	12 a	4 c	9 g	7 t		
ORIGIN	/note="Primer"					
Query Match	0.5%	Score 24;	DB 6;	Length 32;		
Best Local Similarity	100.0%;	Pred. No.	0.39;			
Matches	24; Conservative	0; Mismatches	0; Indels	0; Gaps 0;		
Qy	3780	aatgtgacgcgtgaataataaacgg	3803			
Dd	9	AATGTGACGGTGAATAATACGG	32			
RESULT	6					
LOCUS	AF240004	Sesamum indicum llis globulin mRNA,	1693 bp	mRNA linear PLN 02-MAR-2001		
DEFINITION	AF240004					
ACCESSION	AF240004					
VERSION	AF240004.1	GI:13183172				
KEYWORDS	.					
SOURCE	sesame.					
ORGANISM	Sesamum indicum					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
	Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.					
REFERENCE	1 (bases 1 to 1693)					
AUTHORS	Tai,S.S.K. and Tzen,J.T.C.					
TITLE	Molecular cloning of three storage proteins in sesame					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 1693)					
AUTHORS	Tai,S.S.K. and Tzen,J.T.C.					
TITLE	Direct Submission					
JOURNAL	Submitted (01-MAR-2000) Graduate Institute of Agricultural Biotechnology, National Chung-Hsing University, Taichung 40227,					

```

FEATURES
source
    location/Qualifiers
    1..1683
    /organism="Sesamum indicum"
    /db_xref="taxon:4182"
CDS
    24..1517
    /note="seed storage protein; storage protein in seed

```

```
protein_bodies"
/codon_start=1
/product="IIS globulin"
/protein_id="AAK15087.1"
/db_xref="GI:13183173"
/translation="MALTSLSGFIVVTLLIRGLSAQLAGEQFYWDQLQSOQOQHLKQ
APNCPQVET TACPEPTTDEEACITETPDNNQCEKCAUAVENYICDGLLIDPV
```

BASE COUNT 442 a

Query Match 0.5%; Score 24; DB 8; Length 1683;
 Best Local Similarity 100.0%; Pred. No. 0.39;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3555 ttcaagaccacgacaaacgcgatg 3578
|||||
Db 1263 TTCAAGACCAACGACACGGCGATG 1286

RESULT	7				
AC034196					
LOCUS	AC034196	133108 bp	DNA	linear	HTG 29-MAY-2000
DEFINITION					


```

misc_feature /note="assembly_name:Contig16"
2851. 3964
misc_feature /note="assembly_name:Contig17"
4065. 5164
misc_feature /note="assembly_name:Contig20"
5265. 6462
misc_feature /note="assembly_name:Contig23"
6563. 7875
misc_feature /note="assembly_name:Contig26"
7976. 9431
misc_feature /note="assembly_name:Contig28"
9532. 11077
misc_feature /note="assembly_name:Contig29"
11178. 12990
misc_feature /note="assembly_name:Contig30
clone_end:T7
vector_side:right"
misc_feature 13091. 15234
/note="assembly_name:Contig32"
15335. 17069
misc_feature /note="assembly_name:Contig33"
17170. 18806
misc_feature /note="assembly_name:Contig34"
18907. 20498
misc_feature /note="assembly_name:Contig35"
20599. 23090
misc_feature /note="assembly_name:Contig36"
23191. 26302
misc_feature /note="assembly_name:Contig37"
26403. 28833
misc_feature /note="assembly_name:Contig38"
28934. 31329
misc_feature /note="assembly_name:Contig39"
31430. 33222
misc_feature /note="assembly_name:Contig40"
33323. 35570
misc_feature /note="assembly_name:Contig41"
35671. 37854
misc_feature /note="assembly_name:Contig42"
37955. 39204
misc_feature /note="assembly_name:Contig43"
39305. 42732
misc_feature /note="assembly_name:Contig44"
42833. 45249
misc_feature /note="assembly_name:Contig45"
45350. 48684
misc_feature /note="assembly_name:Contig46"
48785. 52145
misc_feature /note="assembly_name:Contig47"
52246. 55937
misc_feature /note="assembly_name:Contig48"
56038. 59494

```

Query Match 0.5%; Score 24; DB 2; Length 133108;

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3803 gtaaaatattgtaataataataa 3826

|||||

Db 120755 GTAAATATGCTAATAATAATAA 120778

RESULT 8

AC022096

LOCUS

DEFINITION

AC022096

20 ordered pieces.

ACCESSION

AC022096.4

GI:9256401

VERSION

HTG; HTGS_PHASE2; HTGS_DRAFT.

KEYWORDS

human.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AC022096 Homo sapiens chromosome 5 clone CTB-43D14, WORKING DRAFT SEQUENCE, linear HTG 18-JUL-2000

20 ordered pieces.

ACCESSION

AC022096.4

GI:9256401

VERSION

HTG; HTGS_PHASE2; HTGS_DRAFT.

KEYWORDS

human.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 148598)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 148598)
DOE Joint Genome Institute.
Direct Submission
Submitted (26-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced gi:7711677.

-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information

Center Project Name: 80802
Center clone name: C1F978SKB_43D14

Summary Statistics

Consensus quality: 133423 bases at least Q40
Consensus quality: 144004 bases at least Q30
Consensus quality: 145802 bases at least Q20
Estimated insert size: 176000; pulse field gel estimation
Estimated insert size: 147698; sum-of-contigs estimation
Quality coverage: 5.34 in Q20 bases; pulse field gel estimation
Quality coverage: 6.36 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 5195: contig of 5195 bp in length
* 5196 5295: gap of unknown length
* 5296 15425: contig of 10130 bp in length
* 15426 15525: gap of unknown length
* 15526 20947: contig of 5422 bp in length
* 20948 21047: gap of unknown length
* 21048 26789: contig of 5742 bp in length
* 26790 26889: gap of unknown length
* 26890 32943: contig of 6054 bp in length
* 32944 33043: gap of unknown length
* 33044 38651: contig of 5608 bp in length
* 38652 38751: gap of unknown length
* 38752 64412: contig of 25661 bp in length
* 64413 79245: gap of unknown length
* 79246 79345: gap of unknown length
* 79346 82132: contig of 2787 bp in length
* 82133 82232: gap of unknown length
* 82233 86478: contig of 4246 bp in length
* 86479 86578: gap of unknown length
* 86579 90700: contig of 4122 bp in length
* 90701 90800: gap of unknown length
* 90801 93106: contig of 2306 bp in length
* 93107 93206: gap of unknown length
* 93207 93356: contig of 6150 bp in length
* 93357 99456: gap of unknown length
* 99457 116776: contig of 17320 bp in length
* 116777 116876: gap of unknown length
* 116877 117926: contig of 1050 bp in length
* 117927 118026: gap of unknown length
* 118027 127710: contig of 9684 bp in length
* 127711 127810: gap of unknown length
* 127811 136755: contig of 8945 bp in length
* 136756 136855: gap of unknown length
* 136856 138975: contig of 2120 bp in length
* 138976 139075: gap of unknown length

* 139076 146683: contig of 7608 bp in length
 * 146684 146783: gap of unknown length
 * 146784 148598: contig of 1815 bp in length.

FEATURES

Location/Qualifiers
 1..148598
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone_lib="Caltech human BAC library B"

BASE COUNT 41000 a 35188 c 33298 g 37210 t 1902 others
 ORIGIN

Query Match 0.5%; Score 24; DB 2; Length 148598;

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3955 catgaattttataattttataaaat 3978

|||||

Db 124720 CATGAATTTAAATTTTATAAAAT 124743

RESULT

AC026193/C

LOCUS AC026193 152602 bp DNA linear HTC 29-MAY-2000
 DEFINITION Homo sapiens chromosome 3 clone RP11-416H8 map 3p, WORKING DRAFT
 SEQUENCE, 57 unordered pieces.

ACCESSION AC026193

VERSION AC026193.2 GI:8101213

KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 152602)

AUTHORS Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H.,
 Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, G., Li, J., Li, L.,
 Li, S., Li, X., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y.,
 Luo, J., Niu, Y., Qi, Q., Qi, X., Song, S., Sun, M., Sun, W., Sun, Y.,
 Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R.,
 Wang, X., Wang, X., Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y.,
 Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H.,
 Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y.,
 Zhang, Z., Zhu, B., Yu, J. and Yang, H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 152602)

AUTHORS Sun, W., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B.,
 Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y.,
 Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y.,
 Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D.,
 Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J.
 and Yang, H.

Direct Submission

JOURNAL Submitted (21-MAR-2000) Human Genomic Center, Institute of
 Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
 100101, P.R.China

COMMENT

On May 29, 2000 this sequence version replaced gi:7272010.

-----Genome Center

Center:Beijing Center

Center code:Beijing

Website: <http://hgsc.igtp.ac.cn>

<http://www.genomics.org.cn>

Contact: hgsc@igtp.ac.cn

----- Project Information

Center project name: 1% project

Center clone name: RP11-416H8

----- Summary Statistics

Sequencing vector: pUC18; 100% of reads

Chemistry: Dye-terminator: ET 55% of reads

Chemistry: Dye-terminator: Big Dye; 45% of reads

Assembly program: Phrap; version 0.990329
 Consensus quality: 115486 bases at least Q40
 Consensus quality: 144075 bases at least Q30
 Consensus quality: 163107 bases at least Q20
 Insert size: 88668; sum-of-contigs
 Quality coverage: 3.78x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 57 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1540: contig of 1540 bp in length
 * 1541 1640: gap of unknown length
 * 1641 2835: contig of 1195 bp in length
 * 2836 2935: gap of unknown length
 * 2936 4034: contig of 1099 bp in length
 * 4035 4134: gap of unknown length
 * 4135 5760: contig of 1626 bp in length
 * 5761 5860: gap of unknown length
 * 5861 7021: contig of 1161 bp in length
 * 7022 7121: gap of unknown length
 * 7122 8431: contig of 1310 bp in length
 * 8432 8532: gap of unknown length
 * 8533 10150: contig of 1619 bp in length
 * 10151 11827: contig of 1577 bp in length
 * 11828 11928: gap of unknown length
 * 11929 13240: contig of 1313 bp in length
 * 13241 13340: gap of unknown length
 * 13341 14802: contig of 1462 bp in length
 * 14803 14902: gap of unknown length
 * 14903 16661: contig of 1759 bp in length
 * 16662 16761: gap of unknown length
 * 16762 18070: contig of 1309 bp in length
 * 18071 18170: gap of unknown length
 * 18171 19601: contig of 1431 bp in length
 * 19602 19701: gap of unknown length
 * 19702 21364: contig of 1663 bp in length
 * 21365 21464: gap of unknown length
 * 21465 23376: contig of 1912 bp in length
 * 23377 23476: gap of unknown length
 * 23477 23479: contig of 2003 bp in length
 * 23480 25579: gap of unknown length
 * 25580 27646: contig of 2067 bp in length
 * 27647 27746: gap of unknown length
 * 27747 29548: contig of 1802 bp in length
 * 29549 29648: gap of unknown length
 * 29649 31701: contig of 2053 bp in length
 * 31702 31801: gap of unknown length
 * 31802 33144: contig of 1343 bp in length
 * 33145 34245: gap of unknown length
 * 34246 34535: contig of 1291 bp in length
 * 34536 34635: gap of unknown length
 * 34636 35728: contig of 1093 bp in length
 * 35729 35828: gap of unknown length
 * 35829 37639: contig of 1811 bp in length
 * 37640 37739: gap of unknown length
 * 37740 40142: contig of 2403 bp in length
 * 40143 40242: gap of unknown length
 * 40243 41571: contig of 1328 bp in length
 * 41572 41670: gap of unknown length
 * 41671 43259: contig of 1589 bp in length
 * 43260 43359: gap of unknown length
 * 43360 45147: contig of 1788 bp in length
 * 45148 45247: gap of unknown length
 * 45248 46975: contig of 1728 bp in length
 * 46976 47075: gap of unknown length
 * 47076 49136: contig of 2061 bp in length
 * 49137 49236: gap of unknown length

```

* 49237 52178: contig of 2942 bp in length
* 52179 52278: gap of unknown length
* 52279 54957: contig of 2679 bp in length
* 54958 55057: gap of unknown length
* 55058 57417: contig of 2360 bp in length
* 57418 57517: gap of unknown length
* 57517 60195: contig of 2678 bp in length
* 60196 60295: gap of unknown length
* 60296 62295: contig of 2000 bp in length
* 62296 62395: gap of unknown length
* 62396 64401: contig of 2006 bp in length
* 64402 64501: gap of unknown length
* 64502 67886: contig of 3385 bp in length
* 67887 67986: gap of unknown length
* 67987 70888: contig of 2902 bp in length
* 70889 70988: gap of unknown length
* 70989 73943: contig of 2955 bp in length
* 73944 74043: gap of unknown length
* 74044 76750: contig of 2707 bp in length
* 76751 76850: gap of unknown length
* 76851 79324: contig of 2474 bp in length
* 79325 79424: gap of unknown length
* 79425 81198: contig of 1774 bp in length
* 81199 81298: gap of unknown length
* 81299 84892: contig of 3594 bp in length
* 84893 84992: gap of unknown length
* 84993 87653: contig of 2861 bp in length
* 87654 87753: gap of unknown length
* 87754 91151: contig of 3398 bp in length
* 91152 91251: gap of unknown length
* 91252 93951: contig of 2700 bp in length
* 93952 94051: gap of unknown length
* 94052 98226: contig of 4175 bp in length
* 98227 98326: gap of unknown length
* 98327 102109: contig of 3783 bp in length
* 102110 102209: gap of unknown length
* 102210 105933: contig of 3724 bp in length
* 105934 106033: gap of unknown length
* 106034 109944: contig of 3811 bp in length
* 109945 109944: gap of unknown length
* 109945 112749: contig of 2805 bp in length
* 112750 112849: gap of unknown length
* 112850 117845: contig of 4996 bp in length
* 117846 117945: gap of unknown length
* 117946 122444: contig of 4499 bp in length
* 122445 122544: gap of unknown length
* 122545 128723: contig of 6179 bp in length
* 128724 128823: gap of unknown length
* 128824 134603: contig of 5780 bp in length
* 134604 134703: gap of unknown length
* 134704 140178: contig of 5475 bp in length
* 140179 140278: gap of unknown length
* 140279 147021: contig of 6743 bp in length
* 147022 147121: gap of unknown length
* 147122 152602: contig of 5481 bp in length.

FEATURES
    source
        Location/Qualifiers
            1..152602
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="3"
                /map="3p"
                /clone="RP11-416H8"
            1..1540
                /note="assembly_name:Contig42"
            1641..2835
                /note="assembly_name:Contig45"
            2936..4034
                /note="assembly_name:Contig46"
            4135..5760
                /note="assembly_name:Contig51"
            5861..7021
                /note="assembly_name:Contig52"
            7122..8431

```

```

/misc_feature
8532..10150
/note="assembly_name:Contig57"

/misc_feature
10251..11827
/note="assembly_name:Contig62"

/misc_feature
11928..13240
/note="assembly_name:Contig63"

Query Match      0.5%  Score 24;  DB 2;  Length 152602;
Best Local Similarity 100.0%;  Pred. No. 0.4;
Matches 24;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY 3803 gtaaaatatatgtaataataaa 3826
      |||
Db 148305 GTAAATATATGTAATAATAATA 148282

RESULT 10
LOCUS AC024377 162433 bp DNA linear HTG 26-MAY-2000
DEFINITION Homo sapiens clone RP11-611H22, WORKING DRAFT SEQUENCE, 24
unorderd pieces.
ACCESSION AC024377
VERSION AC024377.3 GI:8076825
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 162433)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-611H22
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 162433)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,N., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepe,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,
Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hages,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Laroque,K., Lehoczy,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M.,
Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
Travers,M., Triglio,J., Vassiliev,H., Viel,R., Vo.A., Wilson,B.,
Wu,X., Wymann,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.
Direct Submission
Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7321591.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L6258
Center clone name: 611_H22
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731

```


REFERENCE
AUTHORS

1 (bases 1 to 163272)
 Bao,J., Hu,S., Dong,W., Zhang,X., Wang,J., Zhang,Y., Zhang,H.,
 Liu,B., Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y.,
 Niu,Y., Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H.,
 Liu,Y., Li,G., Li,C., Bao,Q., Wang,X., Song,L., Zhang,L.,
 Guo,D., Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L.,
 Feng,X., Yu,J. and Yang,H.

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
 Chromosome 3p genomic sequence
 2 (bases 1 to 163272)
 Li,T., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B., Bao,W.,
 Sun,Y., Wu,Q., Wang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
 Wang,X., Zhang,H., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H.,
 Li,G., Li,C., Bao,Q., Wang,X., Song,L., Zhang,L., Guo,D.,
 Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
 and Yang,H.

TITLE
JOURNAL

Direct Submission
 Submitted (13-DEC-1999) Human Genomic Center, Institute of
 Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
 100101, P.R.China

COMMENT

On Jan 9, 2000 this sequence version replaced gi:6563492.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

1 4841: contig of 4841 bp in length
 * 1 gap of unknown length
 * 4842 26723: contig of 21882 bp in length
 * 26724 57119: gap of unknown length
 * 57120 104873: contig of 30396 bp in length
 * 104874 163272: contig of 47754 bp in length
 * Location/Qualifiers
 1..163272
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3p"
 /clone="RP11-210B17"

FEATURES
source

BASE COUNT 51500 a 30337 c 29864 g 51563 t 8 others
 ORIGIN

Query Match 0.5%; Score 24; DB 2; Length 163272;
 Best Local Similarity 100.0%; Pred. No. 0.4;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3803 gtaaaatattgttaataataa 3826
 Db 84958 GTAAATATATGTATATATATAA 84935
 |||||

RESULT 12

AC026190
 LOCUS AC026190 165077 bp DNA linear HTG 29-MAY-2000
 DEFINITION Homo sapiens chromosome 3 clone RP11-372K9 map 3p, WORKING DRAFT
 ACCESSION AC026190
 VERSION AC026190.2 GI:8101204
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

1 (bases 1 to 165077)
 Bao,J., Bao,Q., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H.,
 Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
 Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L.,

Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y.,
 Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
 Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R.,
 Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
 Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
 Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
 Zhang,Z., Zhu,B., Yu,J. and Yang,H.

Unpublished
 Chromosome 3p genomic sequence
 2 (bases 1 to 165077)
 Sun,M., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
 Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
 Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
 Li,G., Li,C., Bao,Q., Wang,X., Song,L., Zhang,L., Guo,D.,
 Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
 and Yang,H.

Direct Submission
 Submitted (21-MAR-2000) Human Genomic Center, Institute of
 Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
 100101, P.R.China

On May 29, 2000 this sequence version replaced gi:7272007.

COMMENT

-----Genome Center
 Center:Beijing Center
 Website:http://hgci.gtp.ac.cn
 http://www.genomics.org.cn
 Contact:hgcclgtp.ac.cn
 ----- Project Information
 Center project name:1% project
 Center clone name: RP11-372K9
 ----- Summary Statistics

Sequencing vector: pUC18; 100% of reads
 Chemistry: Dye-terminator; ET 55% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 151327 bases at least Q40
 Consensus quality: 159291 bases at least Q30
 Consensus quality: 164481 bases at least Q20
 Insert size: 158591; sum-of-contigs
 Quality coverage: 3.38x in Q20 bases;sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 19 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

1 1161: contig of 1161 bp in length
 * 1162 1261: gap of unknown length
 * 1262 2361: contig of 1100 bp in length
 * 2362 2461: gap of unknown length
 * 2462 4676: contig of 2215 bp in length
 * 4677 4776: gap of unknown length
 * 4777 7509: contig of 2733 bp in length
 * 7510 7609: gap of unknown length
 * 7610 10791: contig of 3182 bp in length
 * 10792 10891: gap of unknown length
 * 10892 15192: contig of 4301 bp in length
 * 15193 15292: gap of unknown length
 * 15293 20869: contig of 5577 bp in length
 * 20870 20963: gap of unknown length
 * 20970 27390: contig of 6421 bp in length
 * 27391 27490: gap of unknown length
 * 27491 35391: contig of 7901 bp in length
 * 35392 35491: gap of unknown length
 * 35492 42412: contig of 6921 bp in length
 * 42413 42512: gap of unknown length
 * 42513 50794: contig of 8282 bp in length
 * 50795 50894: gap of unknown length
 * 50895 57502: contig of 6608 bp in length
 * 57503 57602: gap of unknown length

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L8635

Center clone name: 530_N_7

FEATURES

source

Location/Qualifiers

1..169770

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="17"

/map="17"

/clone="RP11-530N7"

/clone_lib="RP11-11 Human Male BAC"

complement(1..321)

/rpt_family="MER66A"

complement(322..630)

/rpt_family="AluSx"

complement(631..670)

/rpt_family="MER66A"

749..773

/rpt_family="AT-rich"

complement(867..1176)

/rpt_family="AluJo"

complement(1437..1492)

/rpt_family="MIR"

1715..2085

/rpt_family="MLTK"

complement(2140..2448)

/rpt_family="AluSx"

2456..2524

/rpt_family="MLTK"

complement(2709..3019)

/rpt_family="AluSp"

3274..3324

/rpt_family="(TTA)n"

complement(3377..3784)

/rpt_family="MSTB"

complement(3945..4235)

/rpt_family="AluSx"

complement(4565..4630)

/rpt_family="LIMC3"

complement(5215..5528)

/rpt_family="LIME1"

5546..5727

/rpt_family="AluSg"

5736..6045

/rpt_family="AluSx"

6046..6092

/rpt_family="LIPALC"

6107..6448

/rpt_family="AluSx"

6449..6615

/rpt_family="Tigger3(Golem)"

complement(6616..6917)

/rpt_family="AluSg"

6918..7108

/rpt_family="Tigger3(Golem)"

complement(7109..7404)

/rpt_family="AluY"

7405..7686

/rpt_family="Tigger3(Golem)"

complement(7688..8054)

/rpt_family="LIMDA"

8055..8078

/rpt_family="(TTTTA)n"

complement(8079..8316)

/rpt_family="LIMDA"

complement(8318..8627)

/rpt_family="AluSx"

complement(8647..9092)

repeat_region
complement(9094..9916)
/rpt_family="LIME1"
9917..10177
/rpt_family="AluSg"
complement(10178..10672)
/rpt_family="LIME1"
complement(10673..10970)
/rpt_family="AluSx"
complement(10971..11044)
/rpt_family="LIME1"
11116..11224
/rpt_family="CT-rich"
11393..11613
/rpt_family="MIR"
11721..12031
/rpt_family="AluSp"
complement(12120..12263)
/rpt_family="LIMAO"
12264..12292
/rpt_family="(ATG)n"
complement(12293..12489)
/rpt_family="LIMAO"
complement(12469..12490)
/rpt_family="LIM1"
complement(12491..12802)
/rpt_family="AluSx"
complement(12803..13403)
/rpt_family="LIM1"
complement(13395..13598)
/rpt_family="LIM1"
complement(13599..13855)
/rpt_family="MER70A"
complement(13856..14134)
/rpt_family="AluSg"
complement(14135..14361)
/rpt_family="MER70A"
complement(14362..15829)
/rpt_family="LIM1"
15830..15850
/rpt_family="(TTTG)n"
complement(15851..16503)
/rpt_family="LIM1"
16504..16802
/rpt_family="AluSg1"
complement(16803..17192)
/rpt_family="LIM1"
complement(17194..17792)
/rpt_family="LIM1"
17793..17822
/rpt_family="(TTTA)n"
complement(17823..18638)
/rpt_family="LIPALC"
complement(18612..18733)
/rpt_family="LIM3c"
complement(18839..19064)
/rpt_family="LIM3c"
complement(19161..19536)
repeat_region

Query Match

0.5% Score 24; DB 9; Length 169770;

Best Local Similarity 100.0%; Pred. No. 0.4;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3819 aataataataaaagccacaaaag 3842

|||||

Db 16959 AATAATAATAAAAGCCACAAAG 16936

RESULT 14

AC040977

LOCUS

DEFINITION Homo sapiens chromosome 17 clone RP11-589p10 map 17, *** SEQUENCING

IN PROGRESS ***, 7 unordered pieces.

173878 bp

DNA linear

HTG 05-SEP-2001

```

AC040977
AC040977.5 GI:15431174
HTG: HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 173878)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckigalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepe,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Melchior,J., Menes,L., Mihova,T., Miranda,C., Mlenka,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Ollivar,J.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye.W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 5, 2001 this sequence version replaced gi:14336616.
All repeats were identified using RepeatMasker:
Snit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9393
Center clone name: 589_P_10
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 28017: contig of 28017 bp in length
* 28018 28117: gap of 100 bp
* 28118 31285: contig of 3168 bp in length
* 31286 31385: gap of 100 bp
* 31386 77861: contig of 46476 bp in length
* 77862 77961: gap of 100 bp
* 77962 91856: contig of 13895 bp in length
* 91857 91956: gap of 100 bp
* 91957 116594: contig of 24638 bp in length
* 116595 116694: gap of 100 bp
* 116695 132759: contig of 16065 bp in length
* 132760 132859: gap of 100 bp
* 132860 173878: contig of 41019 bp in length.
Location/Qualifiers
1. .173878
/organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-589P10"
/clone_lib="RPC1-11 Human Male BAC"
BASE COUNT 49141 a 39705 c 39091 g 45236 t 685 others
ORIGIN
Query Match 0.5%; Score 24; DB 2; Length 173878;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3819 aataataataataaagccacaaag 3842
|||||
Db 86976 AATAATAATAATAAAGCCACAAAG 86999
|||||
RESULT 15
AC034192/c
LOCUS AC034192 176217 bp DNA linear PRI 11-FEB-2001
DEFINITION Homo sapiens chromosome 3 clone RP11-416N8 map 3p, complete
sequence.
ACCESSION AC034192
VERSION AC034192.4 GI:12745080
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 176217)
Xiong,H., Zhou,Y., Dong,H., Lin,W., Chen,B., Zhang,C., Zhang,Y.,
Cai,Z., Ying,H.F., Wang,H., Gu,W., Zhu,G., Tu,Y., Zhang,X., Jia,J.,
Shen,H., Zhang,D., Wu,C., Lu,G., Zhong,M., Jiang,H., Ren,S., Fu,G.,
Chen,Z. and Huang,M.
Chromosome 3p genomic sequence
Unpublished
2 (bases 1 to 176217)
Sun,Y., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B., Bao,W.,
Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y., Li,G.,
Li,C., Bao,Q., Wang,J., Song,L., Zhang,L., Guo,D.,
Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
and Yang,H.
Direct Submission
Submitted (05-APR-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing
100101, P.R.China
3 (bases 1 to 176217)
Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D., Guo,Z.,
He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,F.,
Li.G., Li.J., Li.L., Li.S., Li.T., Liu,Y., Liu,N., Liu,B., Liu,Y.,
Li.W., Li.W., Li.Y., Luo,C., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,J., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,
Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X.,
Yu.B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu.N.,
Yu.J. and Yang,H.
Direct Submission
Submitted (11-FEB-2001) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing
100101, P.R.China
On Feb 11, 2001 this sequence version replaced gi:8101214.
-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hgci.gtc.ac.cn
http://www.genomics.org.cn
Contact:hgci.gtc.ac.cn
----- Project Information
Center project name:il% project

```

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
FEATURES
source

```


Center clone name: RP11-416N8
 ----- Summary Statistics
 Sequencing vector: pUC18; 100% of reads
 Chemistry: Dye-terminator: ET 55% of reads
 Assembly: Dye-terminator Big Dye; 45% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 179317 bases at least Q40
 Consensus quality: 179556 bases at least Q30
 Consensus quality: 179718 bases at least Q20
 Insert size: 176217; sum-of-contigs
 Quality coverage: 10.64x in Q20 bases; sum-of-contigs

FEATURES
 source

Location/Qualifiers
 1. 176217
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3p"

BASE COUNT 55712 a 32095 c 32227 g 56183 t
 ORIGIN

Query Match 0.5%; Score 24; DB 9; Length 176217;
 Best Local Similarity 100.0%; Pred. No. 0.4;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3803 gtaaaatataatgtaataataaa 3826
 |||||
 Db 140840 GTAAATATATGTAATAATAATAA 140817

Search completed: September 19, 2002, 03:16:48
 Job time: 17403 sec